

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 July 2003 (10.07.2003)

PCT

(10) International Publication Number  
**WO 03/056419 A1**

(51) International Patent Classification<sup>7</sup>: **G06F 3/033**,  
G06K 9/22, 11/18

(74) Agent: AWAPATENT AB; Box 5117, SE-200 71 Malmö  
(SE).

(21) International Application Number: PCT/SE02/02388

(22) International Filing Date:  
19 December 2002 (19.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
0104456-9 28 December 2001 (28.12.2001) SE

(71) Applicant (for all designated States except US): ANOTO  
AB [SE/SE]; Scheelvägen 19 C, S-223 70 Lund (SE).

(72) Inventor; and

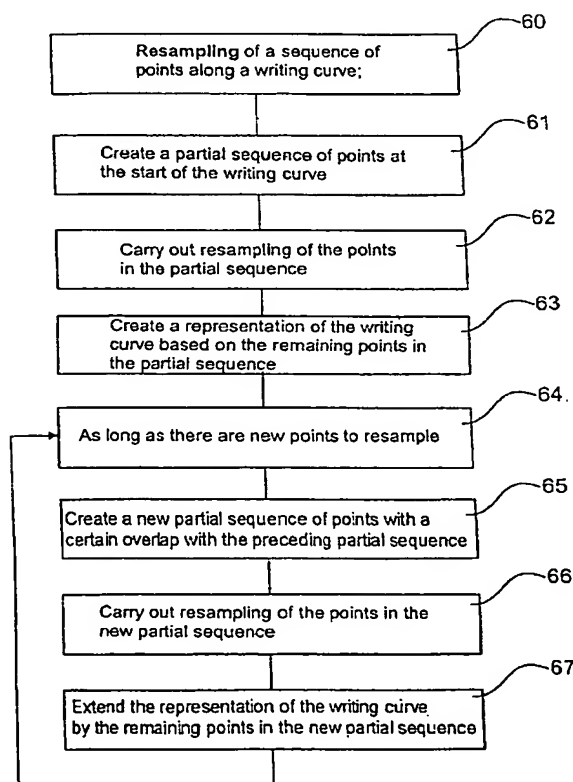
(75) Inventor/Applicant (for US only): BRYBORN, Mattias  
[SE/SE]; Karl XII gata 16 A, S-222 20 Lund (SE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR RECORDING OF ELECTRONIC HANDWRITING



(57) Abstract: According to a method for recording of electronic handwriting, a handwriting movement (1) is digitized into a sequence of points (70) along a writing curve (71) that corresponds to the handwriting movement. The sequence of points is resampled by eliminating those points (89) that are judged not to be necessary for being able to reconstruct the writing curve later with the required visual quality. A resampled representation of the writing curve (71) is created from the points (72, 73, 75) remaining - that is not eliminated - after the resampling. The resampling is carried out on successive and partially overlapping partial sequences (100, 103) of the points (70), the resampled representation of the writing curve (71) being extended successively after resampling of a respective partial sequence. Elected for publication: Fig 6

WO 03/056419 A1

**Declaration under Rule 4.17:**

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent

(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

**Published:**

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD AND APPARATUS FOR RECORDING  
OF ELECTRONIC HANDWRITING

Field of the Invention

The present invention relates to electronic recording of handwritten information. More specifically, the invention relates to recording of electronic handwriting, where a handwriting movement is digitized into a sequence of points along a writing curve that corresponds to the handwriting movement, where the sequence of points is resampled by eliminating those points that are judged not to be necessary for being able to reconstruct the writing curve later with the required visual quality, and where a resampled representation of the writing curve is created from the points remaining - that is the points that have not been eliminated - after the resampling.

Background Art

Today electronic recording of handwritten information has many applications. An example of a writing tool suitable for this purpose is described in WO 01/71473 in the form of an electronic pen. An electronic pen will be used throughout this document to represent a writing tool suitable for recording handwritten information, without the invention in any way being restricted to such a pen.

Electronic pens are intended to be used in hand-held and portable applications, and for this reason, because of consequent restrictions on power consumption, space and cost, it is necessary to store the recorded electronic handwriting as efficiently as possible and in a way that uses as little memory as possible.

US-A-6,101,280 describes a method and a device for compressing electronic handwriting, consisting of

a sequence of pen strokes, each of which constitutes a sequence of pairs of (x, y)-coordinates for points arranged (sampled) in time order along the pen stroke. In order to specify each pair of coordinates with adequate  
5 resolution, nominally one 16-bit integer is required for the x-coordinate and the y-coordinate, respectively. Prior-art technique, including that from US-A-6,101,280, shows, however, a series of measures that can be taken in order to reduce the storage space required for the pen  
10 strokes.

By not storing the points' coordinates as absolute values but instead as difference or delta values relative to the immediately preceding point, a considerable reduction in data is achieved. It is normally sufficient to  
15 have 8 bits per relative coordinate, compared to twice that for absolute coordinates. Even better data reduction is achieved by making use of some form of prediction or extrapolation, for example polynomial approximation, of the expected position of a given point based on a number  
20 of preceding points in the sequence. Each point will thus be represented by the error or deviation between its expected position and its current position.

Before the final storage of the coordinates of the recorded points, the use is known of data-compressing  
25 source coding - often statistical coding such as Huffman coding or arithmetic coding - of the coordinates that have been relative coded in accordance with the above.

Even before the relative coding and the data-compressing source coding, US-A-6,101,280 shows, however,  
30 that many of the recorded points can be eliminated without significantly worsening the visual quality when the electronic handwriting is later to be recreated, for example displayed on a screen. For this purpose, so-called resampling of the sequence of recorded points is  
35 carried out. The Douglas-Peucker's algorithm is often used for resampling a discrete curve, for example consisting of a sequence of points along a recorded pen

stroke, and this is carried out in the following way. A curve segment between a first point and a second point, which is located at a distance from the first point, has a straight line drawn experimentally between the first  
5 and the second point. If the maximal distance between this line and any point on the curve segment is less than a limit value, the curve segment is replaced by the straight line. A straight line can be represented by considerably less data than a curve segment, whereby the  
10 saving in storage space is evident. On the other hand, if the maximal distance exceeds the limit value, a first shorter line is created between the first point and a midpoint on the curve segment and a second shorter line is created between this midpoint and the second point,  
15 the above procedure being repeated recursively on these shorter lines. After the resampling has been concluded, only those points that have not been replaced by a line are saved. The curve segment is thus represented by these points remaining after the resampling.

20 Although the resampling procedure in US-A-6,101,280 is fully capable of real data reduction, it is still associated with certain disadvantages. As the resampling is not carried out until a complete pen stroke, from "pen-down" to "pen-up", has been entered by the user, it  
25 can often be a question of long sequences of points and thus large amounts of coordinate data that have to be processed at one and the same time during the resampling. The storage and processing of large amounts of coordinate data require of course the available working memory to be  
30 sufficiently large. It is, however, far from the case that an electronic pen or a similar portable miniaturized handwriting device has a large working memory; on the contrary, the constant market demands for low cost/small size/low power consumption mean that the working memory  
35 of the pen (or the like) often has a limited storage capacity. In order to be able also to carry out resampling of long sequences of points (long pen strokes) it is

thus necessary for the pen to be provided with a working memory that is (over-)dimensioned for the worst case scenario, or it is necessary not to carry out resampling of sequences of points that are too long. Both alternatives have evident disadvantages.

Another disadvantage of the resampling in accordance with the above is that time delays can arise for the resampling, the subsequent point coordinate coding and the storage, for example if the user pauses while making a pen stroke without lifting the pen. Only later, once the pen stroke has been completed and the pen has been lifted from the writing base, can the sampling and the subsequent operations commence.

#### Summary of the Invention

A main object of the invention is therefore to remedy or at least reduce the above-mentioned problems that are associated with resampling a digitized pen stroke or a similar writing curve.

The object of the invention is in particular to make it possible to resample long sequences of points in spite of a limited available working memory, which resampling is in addition to be initiated and implemented quickly, even if there is a pause during the pen stroke.

More specifically, these objects are achieved by means of a method, an apparatus and a computer program product for recording of electronic handwriting according to the independent claims.

A first aspect of the invention is thus a method for recording of electronic handwriting, where a handwriting movement is digitized into a sequence of points along a writing curve that corresponds to the handwriting movement. The sequence of points is resampled by eliminating those points that are judged not to be necessary for being able to reconstruct the writing curve later with the required visual quality. A resampled representation of the writing curve is created from the points remain-

ing - that is not eliminated - after the resampling. The resampling is carried out on successive and partially overlapping partial sequences of the points, and the resampled representation of the writing curve is extended successively after resampling of a respective partial sequence.

The thus obtained resampled representation of the writing curve may be stored as digital data in any storage means, including but not limited to a register in a processor (CPU), a cache memory, a random access memory or a persistent memory, either permanently or temporarily (the latter case including streaming or other transmission of said digital data from one device to another).

The resampling is carried out according to a preferred embodiment using a line-generalization algorithm, such as the Douglas-Peucker's algorithm, for polygonal approximation of the writing curve, the curve segment between two consecutive remaining points being represented by a straight line.

According to a preferred embodiment, where a partial sequence comprises  $n$  points, the overlap between two consecutive partial sequences is  $w$  points, and  $1 < w \leq n/2$ , an individual partial sequence is stored during the resampling of the same in a memory buffer with room for  $n$  points, after which the last  $w$  points in this individual partial sequence are left in the memory buffer and are included in the resampling of a next following partial sequence. After the resampling of said individual partial sequence, the remaining - that is not eliminated - points in the memory buffer are added to the resampled representation of the writing curve, with the exception of the remaining points from a subset of said last  $w$  points in said individual partial sequence, which instead are included in the resampling of a next following partial sequence as regards addition to the resampled representation of the writing curve.

More specifically, the preferred embodiment comprises the steps of

- a) filling the memory buffer with points;
- b) applying the line-generalization algorithm to the  
5 points in the memory buffer, wherein such remaining points are marked which are not eliminated by the algorithm;
- c) letting the marked points from the points 1 to  $n-x$  in the memory buffer be incorporated in the resampled  
10 representation of the writing curve;
- d) deleting the first  $n-w$  points in the memory buffer;
- e) placing the remaining  $w$  points at the start of the memory buffer;
- 15 f) as long as points are available for resampling:
  - f1) filling up the memory buffer with  $n-w$  new points;
  - f2) applying the line-generalization algorithm to the points in the memory buffer, in accordance  
20 with step b);
  - f3) letting the marked points from the points  $w-x$  to  $n-x$  in the memory buffer be incorporated in the resampled representation of the writing curve;
  - f4) deleting the first  $n-w$  points in the memory  
25 buffer; and
  - f5) placing the remaining  $w$  points at the start of the memory buffer;
- and
- g) letting the marked points from the remaining  
30 points  $n-x+1$  to  $n$  in the memory buffer be incorporated in the resampled representation of the writing curve.

Based on the resampled representation of the writing curve, the remaining points can advantageously be relative coded and compressed using statistical coding. This  
35 is advantageously carried out as soon as the handwriting movement has been digitized into a predetermined number of points. Alternatively, the resampled representation



of the writing curve can be stored as a data set in a memory.

As the resampling according to the invention works in stages with partially overlapping sequences of points, and as results from the individual resamplings are compiled as a combined data representation of the recorded and sampled sequence of points, this makes it possible for the resampling to be used also in electronic pens (and the like) with limited working memory.

The invention has in addition an advantage relating to speed, as the resampling may be initiated as soon as a point-coordinate buffer in the working memory has been filled with recorded points to a sufficient extent.

Another advantage of the invention is its ability to avoid or minimize, automatically, artefacts at "knot points" - points that represent a sharp bend or a sudden change of direction in the handwriting, for example in the form of a sharp tip or a little loop with a tight bend. In fact, such points are often particularly important to legibility, and it is thus desirable to prevent them from being unintentionally eliminated during the resampling. "Knot points" are dealt with in the above-mentioned US-A-6 101 280, but in a way that is resource-intensive - "knot points" are explicitly searched out in the sequence of recorded points prior to the resampling which is then based on identified "knot points".

The invention can finally be summarized as a new and innovative method for resampling a sequence of recorded points "on the fly" or in real-time, that is in immediate association with ongoing digitizing of a handwriting movement, the method being carried out using a sliding and partially overlapping window.

A second aspect of the invention is an apparatus for recording of electronic handwriting, comprising digitizing means, resampling means and a memory, said digitizing means being arranged to digitize a handwriting movement into a sequence of points along a writing curve corre-

sponding to the handwriting movement and said resampling means being arranged to eliminate those points in the sequence of points that are judged not to be necessary for being able to reconstruct the writing curve later with the required visual quality, and to create a resampled representation of the writing curve from the points remaining - that is not eliminated - after the resampling. Said resampling means is arranged to carry out the elimination of points for successive and partially overlapping partial sequences of points, and to extend the resampled representation of the writing curve successively after resampling a respective partial sequence.

The control unit of the apparatus can advantageously be arranged to carry out the method according to the first aspect of the invention. It can also advantageously be designed as an electronic pen.

A third aspect of the invention is a computer program product that can be directly read into a memory associated with a processor, comprising program code for carrying out steps according to the first aspect of the invention.

The second and third aspects of the invention have essentially the same advantages as the first aspect.

Other objects, advantages and characteristics of the invention are apparent from the following detailed description of the invention, from the appended claims and from the drawings.

#### Brief Description of the Drawings

The invention will now be described in more detail with reference to the accompanying drawings.

Fig. 1 is a schematic overview of a system for electronic recording of handwriting according to an embodiment, comprising among other things an electronic pen and a server for receiving entered information from the pen.

Fig. 1a illustrates an example of information - in the form of a handwritten piece of text in the Swedish language - that can be recorded using an electronic pen according to Fig. 1.

5 Fig. 2 is a schematic drawing of the electronic pen in Fig. 1.

Fig. 3 is a schematic drawing of a position-coding pattern, which is applied to a writing base for the electronic pen in Fig. 1.

10 Fig. 4 shows a block diagram of some of the components of the electronic pen in Fig. 1.

Fig. 5 shows a general flow chart for processing of electronic handwriting, the present invention pertaining to a resampling phase in this processing.

15 Fig. 6 is a flow chart of the resampling according to the invention.

Figs 7 and 8 illustrate schematically the principle of resampling recorded points along the curve of a writing movement according to the Douglas-Peucker's algorithm.

20 Fig. 9 illustrates a part of a writing curve that contains a "knot point" - that is a point that represents a sharp bend or sudden change of direction in the handwriting.

25 Fig. 10 illustrates how resampling according to the preferred embodiment is carried out on two successive and partially overlapping partial sequences of points (that is two partially overlapping curve segments of a writing curve).

#### Detailed Description of the Invention

30 By way of introduction, general information is given about the different components of the invention. Later follows a more detailed description of the invention.

\* \* \*

35 A system for electronic recording of handwriting is shown in Fig. 1. An electronic pen 10 is utilized in the

system, which pen will be described in greater detail with reference to Fig. 2. When the user moves the pen 10 in the required pen movements 1 across a writing base 2, the pen movements are recorded as a plurality of digital pen strokes, which can be stored locally in the pen awaiting later transmission to a server 3 via a wireless communication link 4. In order to make this recording possible, the writing base 2 is provided with a position-coding pattern 20, which will be described in greater detail with reference to Fig. 3. A possible application, among many others, is that the handwriting entered by means of the pen movements 1 is incorporated in or attached to an e-mail message in the form of an image object of any common type (for example JPEG, GIF, SVG, PNG or TIFF), which is transmitted via a global network (WAN - "Wide Area Network", for example the Internet) to a recipient specified by the user of the pen. Another application can be that the entered handwriting undergoes computerized text interpretation (ICR - "Intelligent Character Recognition") in the server 3 or in some other local or remote device, including but not limited to a personal computer, a mobile telephone or a portable digital assistant, with or without the use of a network for transmitting the entered handwriting. Another application may be that the handwriting is simply displayed on a screen of the server 3, or of said local or remote device.

Irrespective of the application, the handwriting recorded by the pen 10 is generally stored temporarily in the pen, before it is transferred to the server 3, etc., at a given time.

In the following, a short description is given of the general components in the electronic pen 10 with reference to Fig. 2. A more complete description of the pen 10 is given in WO 01/16691, WO 01/26032 and WO 01/26033, which are herewith incorporated in their entirety by reference.

The electronic pen 10 has a casing or a pen body 11 which has approximately the same shape as the casing of a conventional marker pen. One end of the casing has a window 12, through which images are recorded. The casing  
5 11 contains principally an optics part, an electronics part and a power supply.

The optics part comprises at least one illuminating light source 13, a lens system (not shown in the Figure) and an optical image reader 14. The light source 13,  
10 suitably a light-emitting diode, illuminates a part of the base 2 which can be viewed through the window 12, preferably by means of infrared light or alternatively light of some other wavelength. The base 2 is provided with the position-coding pattern 20. An image of the base  
15 2 is projected on the image reader 14 by means of the lens system.

The power supply for the sensor device 10 is advantageously a battery 15, which alternatively can be replaced by or supplemented by mains power (not shown).

20 The electronics part 16 comprises a control unit 16a to which a storage means 16b is connected. The control unit 16a is responsible for the different functions in the electronic pen 10 and can advantageously be implemented by a commercially available microprocessor such  
25 as a CPU ("Central Processing Unit"), by a DSP ("Digital Signal Processor") or by some other programmable logical device, such as an FPGA ("Field Programmable Gate Array") or alternatively an ASIC ("Application-Specific Integrated Circuit"), discrete analog and digital  
30 components, or some combination of the above.

A conventional pen point 17 is arranged on the casing 11. By means of the pen point 17, the user can write or draw physically on the base 2 by an ordinary pigment-based marking ink being deposited on the surface. The  
35 marking ink in the pen point 17 is suitably transparent to infrared light in order to avoid interference with the opto-electronic detection in the electronic pen 10.

The electronics part comprises in addition a combined transmitter and receiver (transceiver) 18 for sending information to and from a remote device, such as a computer or mobile telephone, or the server 3. The  
5 combined transmitter and receiver 18 is advantageously arranged for short-range radio communication in accordance with the Bluetooth standard at 2.4 GHz in the ISM frequency band ("Industrial, Scientific and Medical"). The combined transmitter and receiver can,  
10 however, alternatively be arranged for infrared communication, such as IrDA ("Infrared Data Association"), or for cable-based communication (such as USB or RS232), or essentially for any other available standard for short-range communication between a hand-held device and  
15 a remote device.

Even though the transmission of information in one embodiment takes place directly between the pen 10 and the server 3, it is to be noted that this can just as well take place via an intermediate device, for example a  
20 mobile telephone, a hand-held computer or a portable personal computer. In such a case, the intermediate device is provided with a combined transmitter/ receiver corresponding to the transmitter/receiver 18 in the pen 10, by means of which information can be transmitted from  
25 the pen to the intermediate device. The latter is also provided with a suitable interface for communication with the server 3 - for example a network card (for communication via a local or global network), or alternatively an analog or digital modem (for communication via a cable-based fixed telephone network, a  
30 mobile telephone network or a satellite telephone network). In this way, the information from the pen can be forwarded to the server 3 by this intermediate device.

In addition, the electronics part can comprise buttons 19a, by means of which the user can control the  
35 functions in the electronic pen 10. The electronic pen 10

can also comprise a display 19b, such as a liquid crystal display, and a lamp 19c for indicating status.

The electronic pen 10 is illustrated further in Fig. 4. The pen's storage means 16b consists as shown of two parts, a working memory 5 and a permanent memory 7. A part of the working memory 5 is taken up by a coordinate buffer 6 for temporary storage of a partial sequence of digitized points during resampling. The working memory 5 is also used in a conventional way by the control unit 16a for reading in program instructions prior to execution and for storage of associated program variables, data structures and temporary results.

The permanent memory 7 contains the set of program instructions which together, during execution by the control unit 16a, are capable of carrying out the resampling method according to the preferred embodiment, and also a representation of a resampled writing curve, which is extended successively by results from the resamplings of the individual partial sequences of points. The complete representation of a resampled and, in the preferred embodiment, compressed writing curve can later be transmitted to the server 3 in the way described above, or alternatively can be streamed over successively to the server 3 as soon as a respective partial sequence of points has been resampled and compressed.

The storage means 16b can in addition also comprise other types of memory, for example a read only memory (ROM). The different types of memory can either be realized as physically different memories such as RAM, SRAM, DRAM, EEPROM, flash memory, etc, or the different types or functions of memories can be implemented as a combined physical memory.

With reference to Fig. 3, the position-coding pattern comprises a virtual raster pattern 21, around which a plurality of marks 22 are placed. Each mark represents one of four possible values from 1 to 4. The value of each mark is represented by its actual position 22 in

relation to its nominal position 23, the latter being at the intersection between a horizontal and a vertical line in the raster pattern 21. Thus, each mark 22 can be located in one of four different positions which are  
5 separated from each other in orthogonal directions from the nominal position 23. The distance is suitably not less than about  $1/8$  and not more than about  $1/4$ , preferably about  $1/6$ , of the distance between two opposite raster lines.

10 The distance between the raster lines can, for example, be about 300 micrometers or about 254 micrometers. The latter distance is particularly suitable for printers and scanners, which often have a resolution that is a multiple of 100 dpi (dots per inch).

15 Each mark 22 consists of an essentially circular dot with a radius that is preferably between about 25% and about 120% of the distance between the dots and the nominal position 23. Alternatively, the marks 22 can be other geometric shapes than circular, such as  
20 rectangular, triangular, elliptical, and can in addition be solid or open.

The position-coding pattern 20 can be constructed so that it codes a very large number of absolute positions. For example, 6x6 adjacent marks can in combination code  
25 a position with x- and y-coordinates. By providing the surface on the base 2 with the position-coding pattern 20, an electronic representation can be obtained of the information that is written or drawn on the base using the electronic pen 10 by repeatedly producing images of  
30 the surface when the pen 10 is moved across the surface. In these images, the marks 22 will appear as foreground objects, while the rasters 21 are only virtual and will not appear in the images.

Position-coding patterns of the type outlined  
35 above are described in greater detail in WO 01/16691, WO 01/26032 and WO 01/26033. An alternative position-coding pattern is shown in WO 00/73983. All these docu-



ments are herewith incorporated in their entirety by reference.

\* \* \*

A way of working 50 for processing electronic hand-writing using the electronic pen 10 and the server 3 according to a preferred embodiment of the invention will now be described with reference to Fig. 5. The steps 51-54 are carried out in the pen 10, while the steps 56-58 are carried out in the server 3.

10 In a first step 51, the pen 10 records continuously images of the base 2 (see Fig. 1), when the user moves the pen in at least one writing movement 1. In practice, the user will carry out a large number of writing movements, each representing wholly or partially for example  
15 a letter, a number or a symbol comprised in a piece of text, for example according to Fig. 1a. In order to make this easier to understand, the remainder of the way of working 50 is, however, described with the simplified assumption that only one writing movement 1 is carried  
20 out by the user. Using the position-coding pattern 20, the pen's electronics part 16 determines a sequence of digital points 70 (see Fig. 7) along the writing movement 1. In other words, the physical writing movement 1 is represented electronically by the sequence of digital  
25 points 70, which can be said to create a writing curve 71 corresponding to the writing movement 1.

In a second step 52, a resampling is carried out of the sequence of digital points 70 by certain individual points being eliminated, these points not being necessary  
30 for providing sufficiently good quality in later reconstruction of the writing curve. The resampling according to the invention is illustrated in greater detail in Figs 6-10 and will be described in greater detail later.

Thereafter, further compression of the resampled  
35 sequence of points is carried out in a step 53, which actually consists of two operations - difference coding followed by source coding. None of these operations

constitutes any central part of the invention, but for the sake of completeness a summary is given in the following.

The difference coding according to the preferred  
5 embodiment is carried out in the following way for a sequence of points. By way of introduction, the x- and y-coordinates are obtained for a current point  $p(n)$  that is to be difference coded, expressed in absolute 16-bit values. Thereafter a polynomial  $G(n)$  of order  $N$  is  
10 approximated to the  $N+1$  preceding points  $p(n-1)$ ,  $p(n-2)$ , ...,  $p(n-N-1)$  in the sequence of points. The polynomial approximation is a thoroughly well-known concept within mathematics and is described, for example, in section 7.2 in "Scientific Computing - An Introductory Survey",  
15 Michael T. Heath, McGraw-Hill, ISBN 0-07-115336-5. Tests have shown that polynomials of order 1 or 2 are optimal for the current digitizing speeds (which are often in the range 70-100 Hz). In addition, it has been found that polynomials of order 1 are preferable when the sequence  
20 of points has undergone resampling. As this is the case in the preferred embodiment, this is based on polynomial approximation with first order polynomials; in other words  $N=1$  and a straight line is approximated to the  $N+1 = 1+1 = 2$  immediately preceding points  $p(n-1)$  and  $p(n-2)$   
25 in the sequence of points.

As the points in the sequence of points have known time-related distances which are determined by the digitizing frequency and which, in addition, are indicated in association with the resampling, a predicted point  
30  $p_{pred}(n)$  can be determined by extrapolating the graph (in this case the straight line) for the polynomial  $G(n)$ . In addition, a deviation  $D(n)$  is determined between the actual point  $p(n)$  and the predicted point  $p_{pred}(n)$ .

Then the determined deviation  $D(n)$  is coded using  
35 a data-compression method for statistical source coding such as Huffman coding or arithmetic coding, both of which are well-known principles within the field of

information theory. More specifically, the deviation's x- and y-coordinates,  $D_x(n)$  and  $D_y(n)$  respectively, are coded individually using statistical source coding. In the preferred embodiment, Huffman coding is used, but  
5 other statistical source coding, such as arithmetic coding, can alternatively be used. Huffman coding is described in detail in, for example, "Introduction to Data Compression", second edition, Chapter 3, Khalid Sayood, Morgan Kaufmann Publishers, 2000. Arithmetic  
10 coding is described in Chapter 4 of the same book.

Finally, the compressed deviation  $D_{\text{packed}}(n)$  is stored in the storage means 16b for subsequent transmission, at a given time, via the wireless communication link 4 (Fig. 1) to a secondary memory in the server 3 - in  
15 Fig. 5 represented symbolically by 55.

When the electronic handwriting thus recorded and transmitted to the server 3 is later to be reproduced, the stored, compressed, difference-coded data is first read in a step 56. Thereafter, a decoding of this data is  
20 carried out in a step 57, which essentially consists of the counterpart of operations in the compression step 53. Once the electronic handwriting has been decoded, the current use of the handwriting follows in a step 58, for example in the form of text interpretation of the hand-  
25 writing, visual presentation of the same on a display screen or transmission of the handwriting as an image object in an e-mail message.

\* \* \*

In the preferred embodiment, the resampling is carried out according to the previously described Douglas-Peucker's algorithm. This algorithm is, as has already  
30 been mentioned, well-known - an application is described, for example, in US-A-6,101,280 - however, it is applied here according to the invention in a new and innovative  
35 way. Basic information about the Douglas-Peucker's algorithm can be obtained in "Algorithms for the reduction of the number of points required to represent a

digitized line or its caricature", D. Douglas and T.K. Peucker, 1973, The Canadian Cartographer, Vol. 10, No. 2, pp 112-122. The basic principle for the Douglas-Peucker's algorithm can be summarized briefly in the following way  
5 with reference to Figs 7 and 8:

A curve segment (the writing curve 71), consisting of a sequence of points 70 along a recorded pen stroke 1, has a first point 72 that constitutes the beginning of the curve segment and a second point 73 that constitutes  
10 the end of the curve segment. A straight line 74 is drawn experimentally between the first and the second points. If the maximal distance 76 between this line 74 and any point 75 on the curve segment 71 is less than a limit value  $D_{limit}$ , the curve segment is considered to be able  
15 to be replaced by the straight line. A straight line can be represented by considerably less data than a curve segment, whereby the saving in storage space is obvious. If, on the other hand, the maximal distance 76 exceeds the limit value  $D_{limit}$ , which is the case in Fig. 7, a  
20 first shorter line 87 is created (see Fig. 8) between the start point 72 and an intermediate point on the curve segment (in this case said point 75) and a second shorter line 88 is created between this intermediate point 75 and the end point 73, the above procedure being repeated on  
25 these shorter lines.

As shown in Fig. 8, the maximal distance 86 between the first shorter line 87 and any intermediate point 89 is now less than the limit value  $D_{limit}$ , so all the intermediate points 89 are replaced by the first shorter line  
30 87. In a corresponding way, the second shorter line 88 replaces corresponding points 89 between the point 75 and the end point 73.

After resampling has been completed, only the points 72, 73, 75 that have not been replaced by a line are  
35 saved or marked. The other points 89 are eliminated or left unmarked. The curve segment 71 is thus represented by these points 72, 73, 75 remaining after resampling.

The resampling according to the preferred embodiment will now be further clarified with reference to Figs 6 and 10. The resampling procedure 60 is commenced with a step 61, in which a partial sequence 100 (see Fig. 10) of points in an initial segment of the total writing curve is read into the coordinate buffer 6 in the pen's 5 10 working memory 5 (see Fig. 4). Resampling according to Douglas-Peucker is then carried out in a step 62 for all the points in the partial sequence 100, starting at a point 101 and ending at a point 102. After the 10 resampling, all points that are to be retained are marked in the coordinate buffer 6 in the way described above.

In a step 63, a not yet complete representation of the resampled writing curve is created on the basis of: 15 the first  $n-x$  points from the coordinate buffer 6, where  $n$  is the length of the coordinate buffer 6 and  $x$  is a considerably smaller value than  $n$ , for example a twentieth. More specifically, the representation will consist of the points, from the total  $n-x$  points, that 20 were marked in the resampling step 62. The significance of the parameter  $x$  will be clarified later with reference to Figs 9 and 10. The representation of the sampled writing curve can either be stored in the permanent memory 7 awaiting the whole writing curve having been resampled, 25 or - which is preferable - be stored temporarily in the working memory 5 ready to be supplied as input data for subsequent difference coding and statistical coding (cf. step 53 in Fig. 5). The latter alternative has the advantage that the total processing of the electronic hand- 30 writing takes a shorter time, as the processing of a part of the writing curve can be carried out as soon as a sufficient number of digitized points are available, even if the user pauses during the writing movement 1.

The step 63 is concluded by the first  $n-w$  points in 35 the coordinate buffer 6 being deleted, where  $w$  is larger than  $x$  but considerably smaller than  $n$ . At the same time, the  $w$  remaining points are moved to the start of the

coordinate buffer 6 - either by physical copying/movement of data or by utilizing the coordinate buffer 6 cyclically and letting a pointer indicate where the buffer starts.

5        In step 64, an iterative loop is commenced, which continues as long as there are still incoming (that is newly digitized) points left to process. In a step 65, a new partial sequence of points is created, by the coordinate buffer 6 being filled up until it is full. On  
10   account of the  $w$  remaining points from the previous partial sequence, the new partial sequence - which consists of all the points in the coordinate buffer 6 - will partially overlap the previous partial sequence, more specifically with an overlap of  $w$  points. The new partial  
15   sequence (during the first iteration of the loop 65-67) is designated 103 in Fig. 10 and extends as shown between a start point 104 (located  $w$  points before the end point 102 of the first partial sequence 100) and an end point 105.

20        In a step 66, resampling is then carried out according to Douglas-Peucker for the new partial sequence, that is the points in the coordinate buffer 6. The remaining points are marked as in step 62.

      In a step 67, the representation of the resampled  
25   writing curve created in step 63 is now extended by adding at the end of the representation the remaining/marked points of the points from the coordinate buffer 6, with the exception of marked points from the first  $w-x$  points and the last  $x$  points. This means that the transition  
30   between the new partial sequence 103 and the old partial sequence 100, as regards how the results from the resamplings of these partial sequences are successively added to the representation of the resampled writing curve, will lie at 106 in Fig. 10.

35        In addition, the first  $n-w$  points in the coordinate buffer 6 are deleted and the remaining  $w$  points are moved to the start of the buffer by analogy with step 63. The

21

execution thereafter returns to step 64 for a new iteration of the loop 65-67.

In a concluding step not illustrated in Fig. 6, the representation of the resampled writing curve is completed by adding at the end of the representation marked  
 5 points of the last  $x$  points from the coordinate buffer 6. Thus, a complete, resampled representation of the writing curve is obtained.

It should be noted that all the Douglas-Peucker  
 10 resamplings are carried out on "fresh" points, that is the  $w$  points from the preceding sampling do not retain any marks.

The resampling procedure 60 described above can be implemented by the following pseudo-code:

```

15 -----
Fill buffer with points;      // (length n)

Douglas-Peucker ();          // Mark all points in the buffer that
                              are to be saved
20 StorePoints (1, n-x);      // Save marked points of the first n-x
                              points in the buffer
RemovePoints (n-w);          // Delete first n-w points in the
                              buffer and move the remaining
                              points to the start of the buffer
25
While (Points left)
    Fill buffer with points;
    Douglas-Peucker ();        // Mark all points in the buffer
                              that are to be saved
30    StorePoints (w-x, n-x);  // Save marked points
    RemovePoints (n-w);

End While
StorePoints (n-x+1, n);      // Save the remaining marked points
35 -----

```

The values of  $w$  and  $x$  can be selected to suit the application concerned, however such that

22

$x < w,$   
 $w > 1,$   
 $w \leq n/2.$

n is set in the preferred embodiment to 128. x is  
5 selected as  $w/2$ .

For a sampling frequency of 80 Hz, w should  $\approx n/10$ .  
Experience shows that a relatively larger w is required  
for higher sampling frequencies.

\* \* \*

10 Fig. 9 illustrates an exemplifying scenario, where a  
writing curve for some reason (for example, insufficient  
size of the coordinate buffer) is divided into a first  
partial sequence 91 of points 90 (start point 92, end  
point 93) and a second partial sequence 91' (start point  
15 92', end point 93') with a joint without an overlap at  
95. The second partial sequence 91' contains a "knot  
point" 94, which is located immediately at the start of  
the partial sequence. If the Douglas-Peucker's algorithm  
is applied to these partial sequences 91, 91', there is a  
20 danger of artefacts at the "knot point" 94, that is there  
is a danger that this point is unintentionally eliminated  
during resampling and that information that is valuable  
for legibility is lost.

Due to the fact that the resampling operations  
25 according to the invention are carried out on the par-  
tially overlapping partial sequences 100, 103, the  
sequence of points that are taken into account during  
the resampling is enlarged. In particular, the sequence  
is enlarged so that it also includes the last w points  
30 from the preceding partial sequence, whereby a "knot  
point" 94 is given a "fairer" evaluation and is iden-  
tified as the point worthy of being marked that it  
actually is. Since in the creation of the resampled  
representation of a particular partial sequence, marked  
35 points that lie among the first w-x points and the last x  
points in the coordinate buffer are excluded (cf. step 67  
in Fig. 6), and it could be said that these points are



given to the preceding or subsequent resampling step to deal with, any "knot points" both before and after the transition 106 between consecutive partial sequences 100, 103 are taken into account.

5       For this reason,  $x$  is advantageously selected "symmetrically" in relation to the window  $w$ , that is  $x = w/2$ , so that the same account is taken of points at the immediate commencement of a partial sequence as at the immediate end.

10       The resampling method according to the invention thus reduces, automatically, the danger of an artefact at a "knot point", irrespective of whether it is located at the immediate commencement of a partial sequence (like point 94) or at the end of a partial sequence (that is  
15 just before the end point 93 in Fig. 9). The invention has here an evident advantage in comparison with US-A-6,101,280, in which "knot points" must be located explicitly prior to the resampling.

\* \* \*

20       The sampling procedure 60 described above utilizes overlapping resampling using Douglas-Peucker (steps 62 and 66 respectively, overlap =  $w$  points) but with non-overlapping successive storage of results from the resamplings (steps 63 and 67 respectively), where the  
25 parameter  $x$  indicates 106 the "joint" between successive stored results (see 106 in Fig. 10). In an alternative embodiment, the parameter  $x$  is set to 0, which has the result that both resampling and storage are carried out with the same overlap of  $w$  points between two partial  
30 sequences. Before storage, an OR operation is carried out for this overlapping area, so that an individual point is included during storage if it has been marked during the resampling of at least one of the partial sequences.

35       The invention has been described above in the form of a few exemplifying embodiments. However, the invention is in no way limited to these, but covers many other variants, according to what is defined by the scope of

protection of the appended claims and, in addition, can easily be recognized by a person skilled in the art. For example, other line-generalization or polygonal approximation algorithms or techniques than Douglas-Peucker can  
5 be used for the resampling according to the invention. Some examples of such possible techniques are Jenk's algorithm, Reumann-Witkam's algorithm, so-called "Coordinate-Averaging" or fractal techniques such as "Walking-Divider".

## CLAIMS

1. A method for recording of electronic handwriting, in which

5 a handwriting movement (1) is digitized into a sequence of points (70) along a writing curve (71) that corresponds to the handwriting movement;

the sequence of points is resampled by eliminating those points (89) that are judged not to be necessary for  
10 being able to reconstruct the writing curve later with the required visual quality; and

a resampled representation of the writing curve (71) is created from the points (72, 73, 75) remaining - that is, not eliminated - after the resampling, c h a r a c -  
15 t e r i z e d in that

the resampling is carried out on successive and partially overlapping partial sequences (100, 103) of the points (70), and

the resampled representation of the writing curve  
20 (71) is extended successively after resampling of a respective partial sequence.

2. The method as claimed in claim 1, in which the resampling is carried out according to an algorithm for  
25 polygonal approximation of the writing curve (71), the curve segment between two consecutive remaining points (72, 75) being represented by a straight line (87).

3. The method as claimed in claim 2, in which said  
30 algorithm is the Douglas-Peucker's algorithm.

4. The method as claimed in any one of the preceding claims, in which a partial sequence (100, 103) comprises n points, the overlap between two consecutive partial  
35 sequences is w points, and  $1 < w \leq n/2$ , an individual partial sequence (100) being stored during the resampling in a memory buffer (6) with room for n points, and the

last  $w$  points in this individual partial sequence being left in the memory buffer and being included in the resampling of a next following partial sequence (103).

5           5. The method as claimed in claim 4, in which, after the resampling of said individual partial sequence (100), the remaining - not eliminated - points (72, 73, 75) in the memory buffer (6) are added to the resampled representation of the writing curve (71), with the exception  
10 of the remaining points from a subset of said last  $w$  points in said individual partial sequence, the subset consisting of  $x$  points, which instead are included in the resampling of a next following partial sequence as regards addition to the resampled representation of the  
15 writing curve.

6. The method as claimed in claim 5, comprising the steps of

- a) filling the memory buffer (6) with points;
- 20       b) applying a line-generalization algorithm to the points in the memory buffer, wherein such remaining points (72, 73, 75) are marked which are not eliminated  
by the algorithm;
- c) letting the marked points from the points 1 to  
25  $n-x$  in the memory buffer be incorporated in the resampled representation of the writing curve (71);
- d) deleting the first  $n-w$  points in the memory buffer;
- e) placing the remaining  $w$  points at the start of  
30 the memory buffer;
- f) as long as points are available for resampling:
  - f1) filling up the memory buffer (6) with  $n-w$   
new points;
  - f2) applying the line-generalization algorithm  
35 to the points in the memory buffer, in accordance with step b);

f3) letting the marked points from the points  $w-x$  to  $n-x$  in the memory buffer be incorporated in the resampled representation of the writing curve;

5 f4) deleting the first  $n-w$  points in the memory buffer; and

f5) placing the remaining  $w$  points at the start of the memory buffer;  
and

10 g) letting the marked points from the remaining points  $n-x+1$  to  $n$  in the memory buffer be incorporated in the resampled representation of the writing curve.

7. The method as claimed in claim 5 or 6, in which  $x = w/2$ .

15 8. The method as claimed in any one of claims 5-7, in which  $n = 128$  and  $w \approx n/10$ .

9. The method according to any one of the preceding  
20 claims, comprising the final steps, after resampling, of based on the resampled representation of the writing curve (71), relative coding the coordinates for the points (72, 73, 75) remaining therein; and  
compressing the relative-coded coordinates using  
25 statistical coding.

10. The method as claimed in claim 9, in which the coordinates for an individual remaining point (73) are relative coded by approximating a polynomial of order  $N$   
30 to the  $N+1$  preceding remaining points (72, 75); determining a prediction of said individual point using said polynomial; determining a deviation between said individual point and its prediction; and using this deviation as a relative coding for said individual point.

35 11. The method as claimed in any one of claims 1-10, in which the resampled representation of the writing

curve (71) is stored as a data set in a memory (5, 7, 16b).

12. The method as claimed in any one of the preceding claims, in which the resampling is carried out as soon as the handwriting movement (1) has been digitized into a predetermined number of points (n).

13. The method as claimed in claim 4, in which points (72, 73, 75) remaining - not eliminated - in the memory buffer (6) after the resampling of said individual partial sequence (100) are added to the resampled representation of the writing curve (71), specifically in such a way that the remaining points from the overlapping w points between two consecutive partial sequences (100, 103) are added to the resampled representation of the writing curve, if they have not been eliminated during resampling of at least one of these two consecutive partial sequences.

20

14. An apparatus for recording of electronic handwriting, comprising digitizing means (14, 16a), resampling means (16a) and a memory (16b), said digitizing means (14, 16a) being arranged to digitize a handwriting movement (1) into a sequence of points (70) along a writing curve (71) corresponding to the handwriting movement and said resampling means being arranged to eliminate those points (89) in the sequence of points that are judged not to be necessary for being able to reconstruct the writing curve later with the required visual quality, and to create a resampled representation of the writing curve from the points (72, 73, 75) remaining - that is, not eliminated - after the resampling, characterized in that

35 said resampling means (16a) is arranged to carry out the elimination of points for successive and partially overlapping partial sequences (100, 103) of points (70),

and to extend the resampled representation of the writing curve (71) successively after resampling of a respective partial sequence.

5           15. The apparatus as claimed in claim 14, in which the resampling means (16a) is arranged to carry out the method according to any one of claims 1-13.

10           16. The apparatus as claimed in claim 14 or 15, in the form of an electronic pen (10).

15           17. A computer program product that can be directly read into a memory (16b) associated with a processor (16a), comprising program code for carrying out the steps according to any one of claims 1-13.

18. A method as claimed in claim 6, wherein said line-generalization algorithm is the Douglas-Peucker's algorithm.

1/6

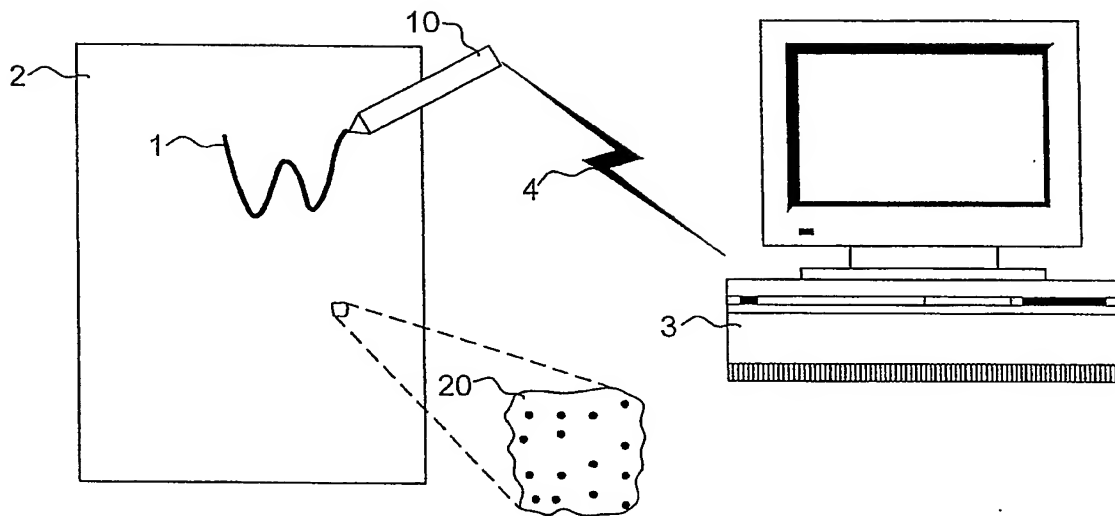


Fig 1

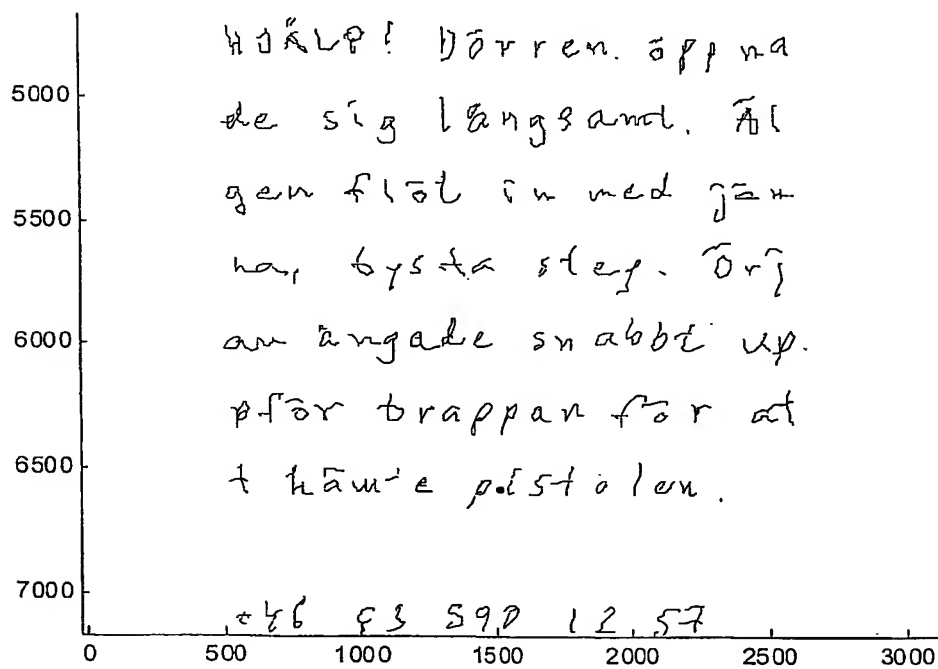
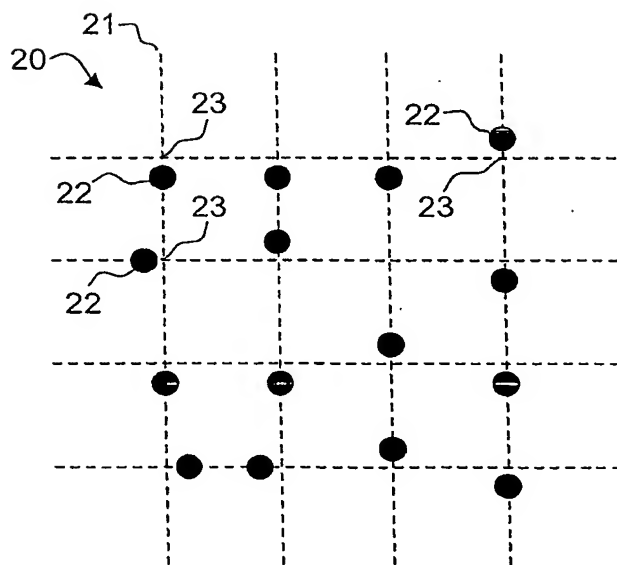
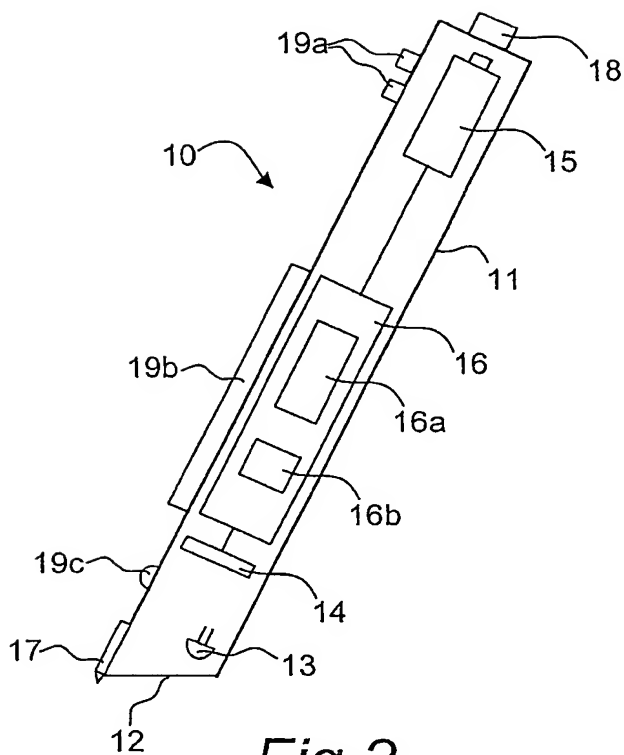


Fig 1a



2/6



3/6

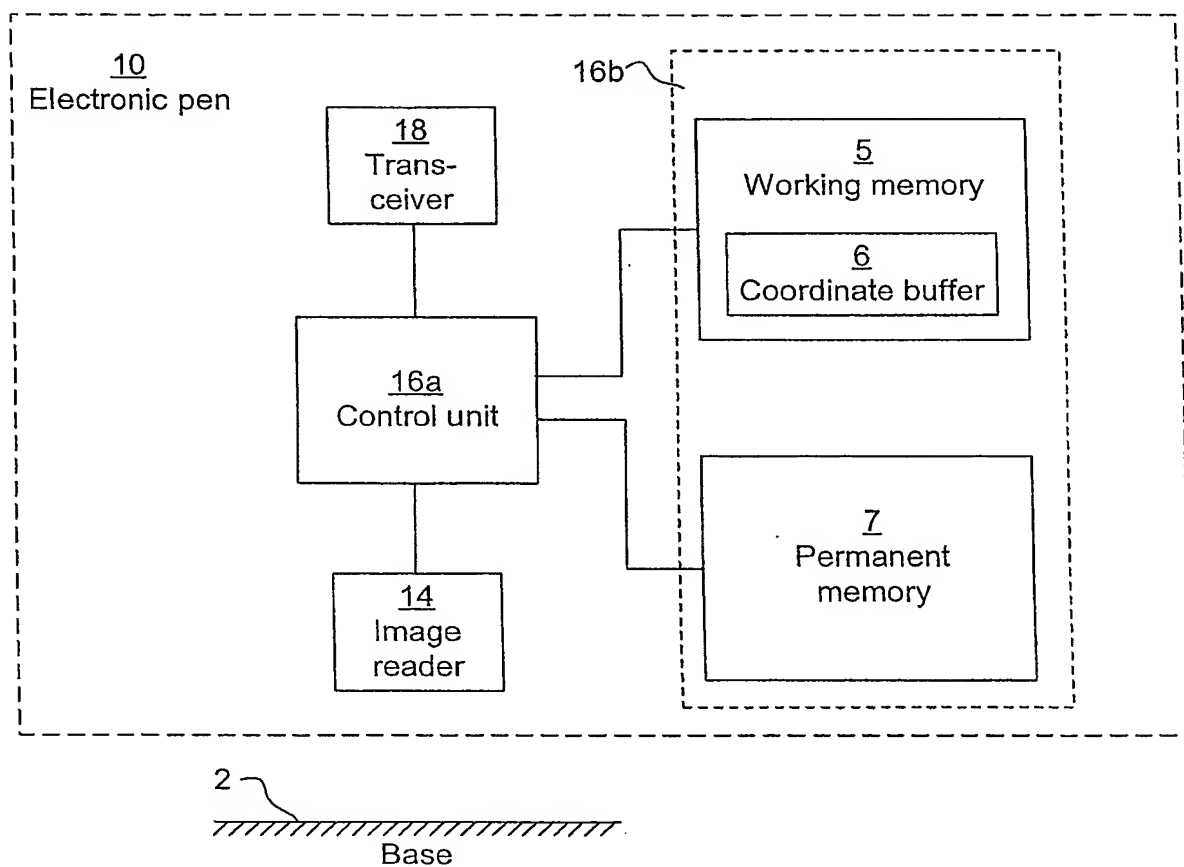
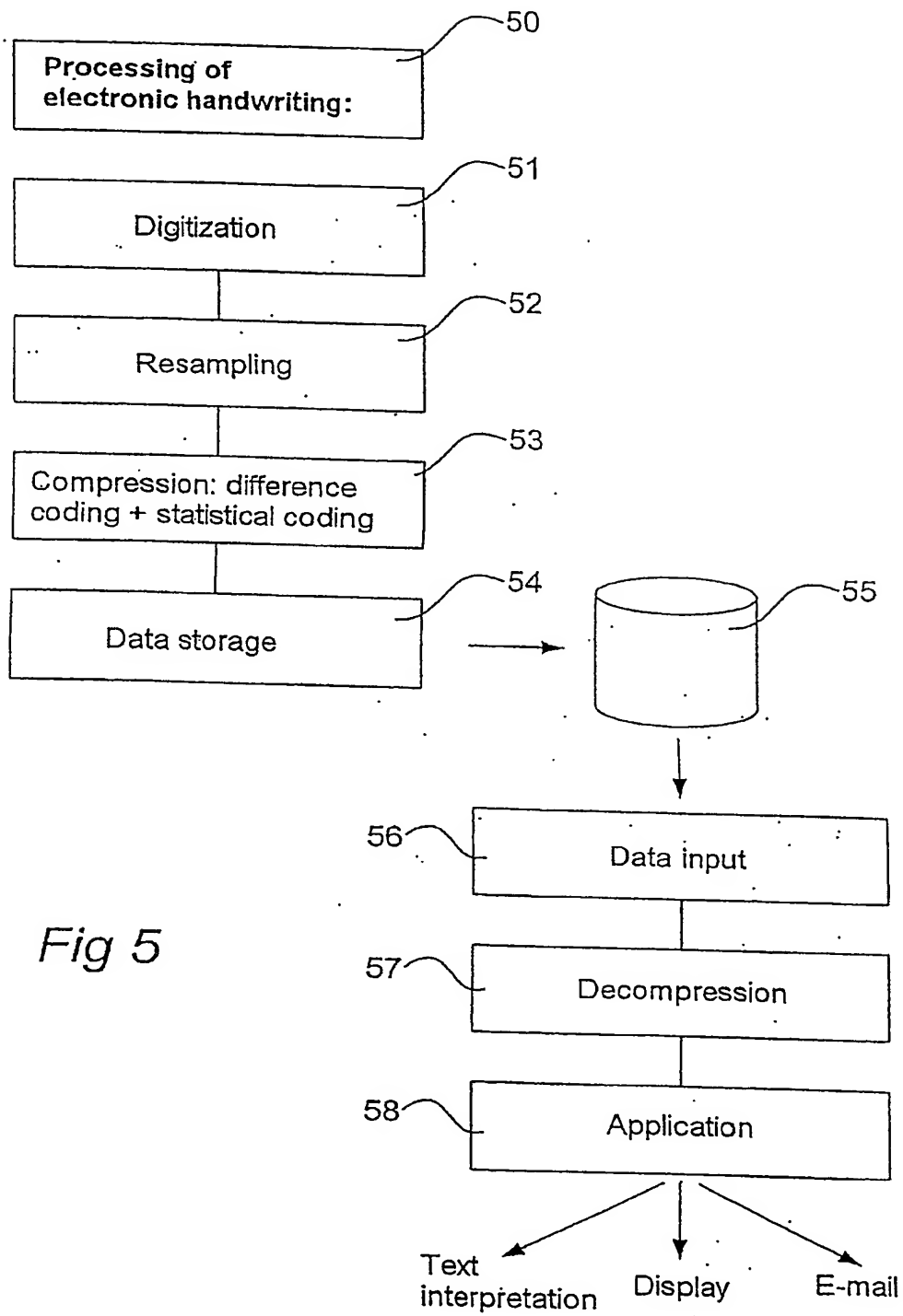


Fig 4

4/6

*Fig 5*

5/6

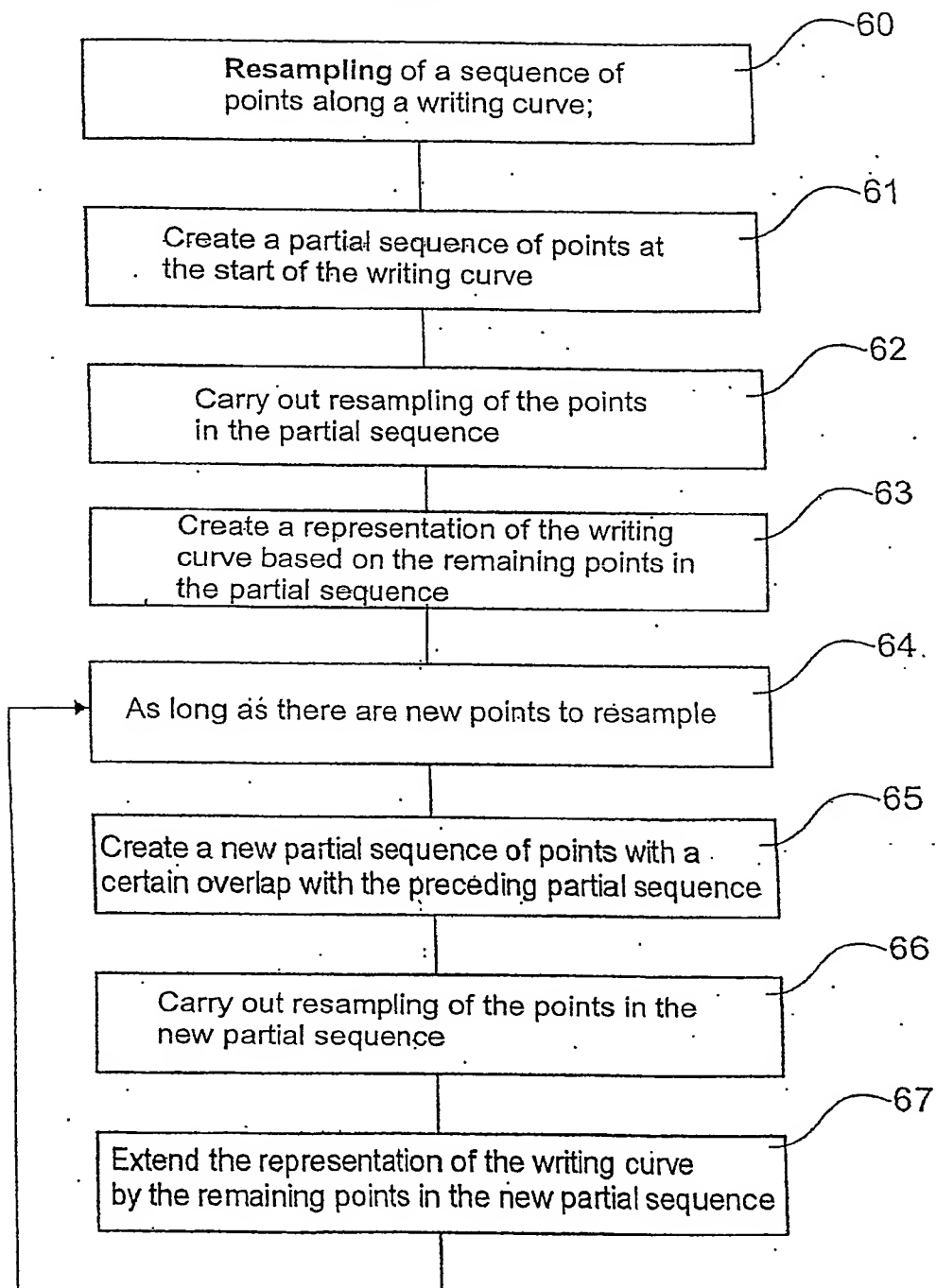


Fig 6

6/6

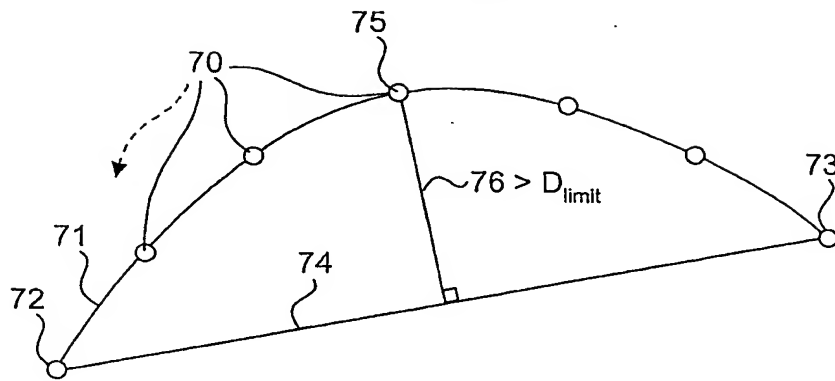


Fig 7

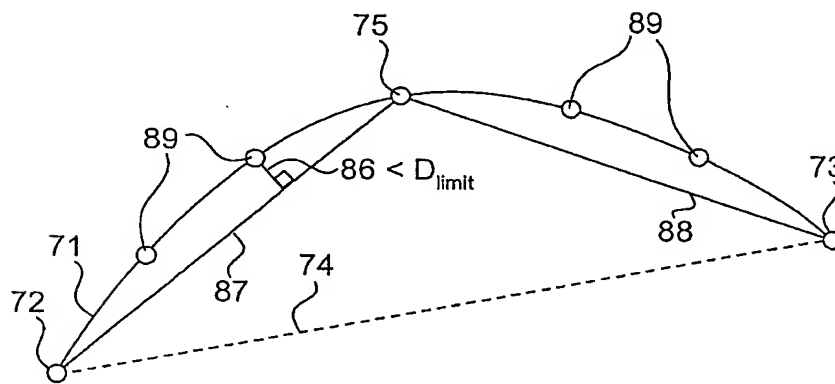


Fig 8

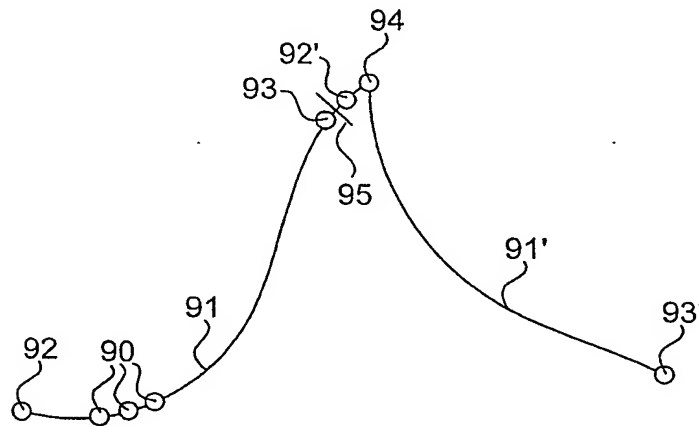


Fig 9

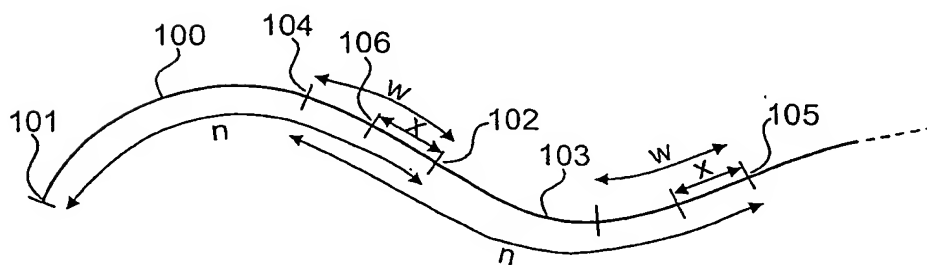


Fig 10

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 02/02388

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: G06F 3/033, G06K 9/22, G06K 11/18  
According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: G06F, G06K, G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4718103 A (SHOJIMA, H. ET AL), 5 January 1988 (05.01.88), abstract --	1-18
A	WO 9403853 A1 (COMMUNICATION INTELLIGENCE CORP), 17 February 1994 (17.02.94), abstract --	1-18
A	EP 0848350 A1 (HEWLETT-PACKARD CO), 17 June 1998 (17.06.98), abstract --	1-18
A	US 6101280 A (REYNOLDS, D.E.), 8 August 2000 (08.08.00), abstract -- -----	1-18

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search	Date of mailing of the international search report
14 March 2003	19 -03- 2003
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer  Åsa Viken /LR Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

Information on patent family members

30/12/02

International application No.

PCT/SE 02/02388

Patent document cited in search report			Publication date	Patent family member(s)		Publication date
US	4718103	A	05/01/88	JP	62082486 A	15/04/87
				KR	9008555 B	24/11/90
-----						
WO	9403853	A1	17/02/94	NONE		
-----						
EP	0848350	A1	17/06/98	CN	1185060 A	17/06/98
				GB	9625661 D	00/00/00
				JP	3091719 B	25/09/00
				JP	10247240 A	14/09/98
-----						
US	6101280	A	08/08/00	CN	1122022 A	08/05/96
				DE	69428527 D,T	08/05/02
				EP	0691623 A,B	10/01/96
				JP	3176023 B	11/06/01
				JP	8044826 A	16/02/96
				US	6018591 A	25/01/00
-----						

**THIS PAGE BLANK (USPTO)**